Odjel za matematiku
Sveučilišni preddiplomski studij matematike

Marinela Knežević

Vigenèreova i Playfairova šifra

Završni rad

Osijek, 2015.

Sveučilište J. J. Strossmayera u Osijeku
Odjel za matematiku
Sveučilišni preddiplomski studij matematike

Marinela Knežević

Vigenèreova i Playfairova šifra

Završni rad

Voditelj: doc. dr. sc. Mirela Jukić Bokun

Osijek, 2015.

Sadržaj

Sažetak

Vigenèreova šifra je metoda šifriranja teksta koja koristi serije različitih Cezarovih šifri na osnovi ključne riječi. U 16. stoljeću je šifru prvi opisao Giovan Battista Bellaso, dok je Blaise de Vigenère osmislio šifru s autoključem. Otkriće metode je pogrešno pripisano Blaise de Vigenèreu u 19. stoljeću i danas je poznata kao "Vigenèreova šifra". Premda je šifra lagana za razumijevanje i implementaciju, kroz tri se stoljeća odupirala svim pokušajima razbijanja, zbog čega je stekla opis "le chiffre indéchiffrable", francuski za neprobojnu šifru. Ipak, pokazalo se da se šifra može razbiti metodama za određivanje duljine riječi i samim određivanjem ključne riječi.
Playfairova šifra je jednostavna, ali učinkovita šifra koja se bazira na tome da se parovi slova šifriraju korištenjem 5×5 matrice. Šifru zapravo nije izumio barun Playfair nego fizičar i izumitelj Charles Wheatstone u 19. stoljeću. Playfairova uloga bila je popularizirati ju. Isprva, šifra je smatrana prekompliciranom te se nije koristila često. No, pokazalo se da je jednostavnija za korištenje od većine drugih šifara koje su onda bile u uporabi. Ova šifra se razbija korištenjem analize frekvencije bigrama.
U radu su definirane Vigenèreova i Playfairova šifra te je na primjerima prikazan način šifriranja i dešifriranja u svakom od ovih kriptosustava.

Ključne riječi: kriptologija, kriptografija, kriptoanaliza, ključ, otvoreni tekst, šifrat, šifriranje, dešifriranje, kriptosustav, Vigenèreova šifra, Kasiskijev test, indeks koincidencije, Playfairova šifra, metoda vjerojatne riječi

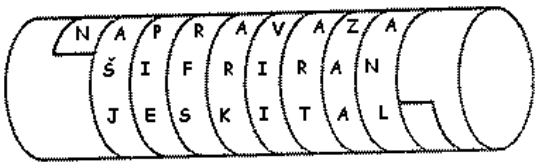
Summary

The Vigenère cipher is a method of encrypting alphabetic text by using series of different Caesar ciphers based on the letters of a keyword. The method was originally described by Giovan Battista Bellaso in 16th century, while Blaise de Vigenère created autokey cipher. The invention was later missattributed to Blaise de Vigenère in the 19th century and is now widely known as the "Vigenère cipher". Though the cipher is easy for understanding and implementation, for three centuries it resisted all attempts to break it; this is why it earned the description "le chiffre indéchiffrable", french for the indecipherable cipher. However, it appeared that the cipher can be broken by using methods for determining length of a keyword and determining a keyword.
The Playfair cipher is a simple but effective cipher in which pairs of letters are encrypted using the 5×5 matrix. The cipher was actually invented not by Baron Playfair, but by the physicist and inventor Sir Charles Wheatstone in 19th century. Playfair's main role was to popularize it. Initially, the cipher was turned down because it was thought to be too complicated. But, it turned out to be easier to use than almost all the ciphers that were then in use. This cipher can be broken by using the analysis of bigrams frequencies.
The Vigenère and the Playfair cipher are defined in this work. Also, methods of encryption and decryption in each of this cryptosystems are shown on examples.

Keywords: cryptology, cryptography, cryptoanalysis, key, plaintext, ciphertext, encryption, decryption, cryptosystem, Vigenère cipher, Kasiski examination, index of coincidence, Playfair cipher, probably word method

Uvod

Kroz cijelu povijest čovječanstva postojala je potreba za sigurnom razmjenom informacija, odnosno potreba za slanjem poruka na način da samo osoba kojoj je ta poruka namijenjena ima mogućnost pročitati sadržaj. Radi toga, korištene su različite metode koje su se s vremenom poboljšavale i usavršavale.
Početak kriptografije nije točno utvrđen, ali se smatra da je počela više od 2000 godina pr. Kr. jer iz tog vremena potječu prvi pronađeni tragovi šifriranja. Točnije, oko 1900. godine pr. Kr. u Egiptu je nastao natpis koji se danas smatra prvim dokumentiranim primjerom pisane kriptografije. Stari Grci su također pridonijeli razvoju kriptografije; Spartanci su u 5. stoljeću pr. Kr. upotrebljavali napravu za šifriranje zvanu *skital*. Skital je bio drveni štap oko kojeg se namotavala vrpca od pergamenta pa se na nju okomito pisala poruka. Nakon upisivanja poruke, vrpca bi se odmotala, a na njoj bi ostali izmiješani znakovi koje je mogao pročitati samo onaj tko je imao štap jednake debljine.



Slika 1: Skital

Kriptografija se s vremenom razvila i usavršila, što je vrlo bitno na globalnim, nacionalnim razinama i slično, ali i u nekim situacijama s kojima se svakodnevno susrećemo, poput slanja obične SMS poruke. Često zapravo nismo niti svjesni koliko nam je bitna zaštita i sigurnost naših osobnih podataka, kao niti što je sve potrebno kako bi to uopće bilo omogućeno.
U ovom radu bavit ćemo se Vigenèreovom šifrom koja je nastala u 16. stoljeću te Playfairovom šifrom koja je nastala u 19. stoljeću. U prvom poglavlju objasnit ćemo ukratko osnovne kriptografske pojmove koje ćemo koristiti. U drugom poglavlju ćemo opisati šifriranje i dešifriranje Vigenèreovom šifrom, dok ćemo šifriranje i dešifriranje Playfairovom šifrom opisati u trećem poglavlju.

Poglavlje 1

Uvod u kriptografiju

U ovom ćemo poglavlju ukratko uvesti i pojasniti pojmove koji će nam trebati u ostatku rada, kao što su osnovni kriptografski pojmovi, shema kriptografskog algoritma te glavne podjele kriptosustava.

1.1 Osnovni pojmovi kriptografije

Kriptologija je znanost koja se bavi izučavanjem i definiranjem metoda za zaštitu informacija i izučavanjem i pronalaženjem metoda za otkrivanje šifriranih informacija (dekriptiranjem). Objekti izučavanja kriptologije su pisane (kriptografija), govorne (kriptofonija), vizualne (slike, karte, sheme) i druge poruke. Rezultate kriptologije prvenstveno koriste oružane snage i diplomatska služba, a razvojem telekomunikacija i mnoge druge službe.
Kriptografija je znanstvena disciplina koja se bavi analizom i pronalaženjem metoda za logičku promjenu podataka. Riječ dolazi od grčkog pridjeva kriptós - "skriven" i glagola gráfo - "pisati". Razvija se radi potrebe da se podatci pošalju primatelju tako da nitko drugi osim primatelja i pošiljatelja ne zna izvorne podatke.
Osnovni pojmovi kojima se služimo su sljedeći:

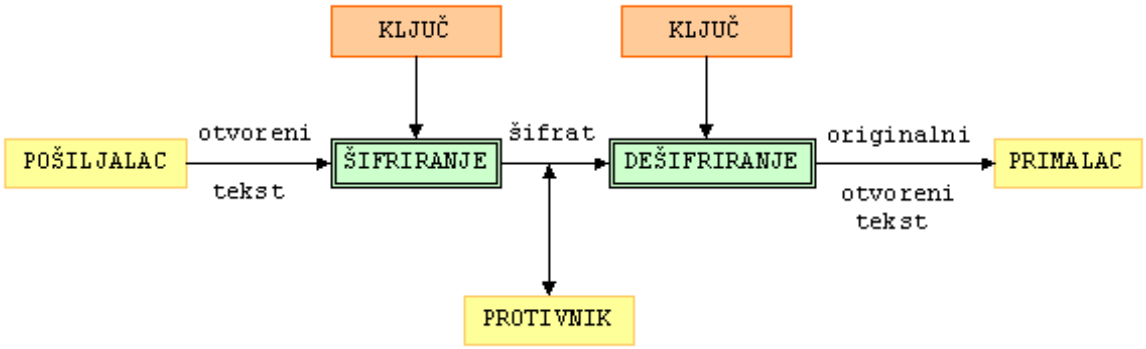
Otvoreni tekst – poruka koju pošiljatelj želi poslati primatelju; to može biti tekst, numerički podatak ili bilo što drugo.

Ključ – način šifriranja i dešifriranja podataka.

Šifriranje (kodiranje) - postupak kojim se otvoreni tekst pomoću ključa promjeni te se više ne može čitati, osim ako osoba posjeduje ključ.

Dešifriranje (dekodiranje) - postupak kojime se šifrirani tekst pomoću ključa vrati u izvorni otvoreni tekst.

U literaturi se pošiljatelj najčešće naziva Alice, primatelj Bob, dok se njihov protivnik, osoba koja pokušava razumijeti njihovu poruku, zove Eve. Osnovni algoritam glasi ovako: pošiljatelj - Alice šifrira otvoreni tekst koristeći unaprijed dogovoreni ključ. Tako šifrirani tekst se naziva **šifrat**. Nakon toga Alice šalje šifrat preko nekog komunikacijskog kanala primatelju - Bobu. Protivnik - Eve prisluškujući može doznati sadržaj šifrata, ali ne može odrediti otvoreni tekst. Za razliku od nje, Bob koji zna ključ kojim je šifrirana poruka može dešifrirati šifrat i odrediti otvoreni tekst.



Slika 1.1: Shema osnovnog algoritma

Kriptoanaliza ili dekriptiranje je znanstvena disciplina koja se bavi proučavanjem postupaka za čitanje skrivenih poruka bez poznavanja ključa.
Kriptologija obuhvaća kriptografiju i kriptoanalizu.

1.2 Kriptografski algoritam ili šifra

Kriptografski algoritam ili šifra je uređeni par dvije funkcije, jedne za šifriranje, a druge za dešifriranje, koje preslikavaju osnovne elemente otvorenog teksta u osnovne elemente šifrata i obratno. Te funkcije se biraju iz određene familije funkcija u ovisnosti o ključu. Skup svih mogućih vrijednosti ključeva nazivamo *prostor ključeva*. Konačno, kriptosustav se sastoji od kriptografskog algoritma te svih mogućih otvorenih tekstova, šifrata i ključeva.

Definicija 1.1. (Kriptosustav) *Kriptosustav je uređena petorka* $(\boldsymbol{P}, \boldsymbol{C}, \boldsymbol{K}, \boldsymbol{E}, \boldsymbol{D})$ *za koju vrijedi:*

1. $\boldsymbol{P}$ *je konačan skup svih mogućih osnovnih elementa otvorenog teksta;*

2. $\boldsymbol{C}$ *je konačan skup svih mogućih osnovnih elemenata šifrata;*

3. ***K*** *je prostor ključeva, tj. konačan skup svih mogućih ključeva;*

4. *Za svaki* $K \in \boldsymbol{K}$ *postoji funkcija šifriranja* $e_K \in \boldsymbol{E}$ *i odgovarajuća funkcija dešifriranja* $d_K \in \boldsymbol{D}$. *Pritom su* $e_K : \boldsymbol{P} \to \boldsymbol{C}$ *i* $d_K : \boldsymbol{C} \to \boldsymbol{P}$ *funkcije sa svojstvom da je* $d_K(e_K(x)) = x$, *za svaki otvoreni tekst* $x \in \boldsymbol{P}$.

Uočimo da iz $d_K(e_K(x)) = x$ u gornjoj definiciji slijedi da funkcije e_K moraju biti injekcije. Naime, ako bi bilo $e_K(x_1) = e_K(x_2) = y$, za dva različita otvorena teksta x_1 i x_2, onda primatelj ne bi mogao odrediti treba li y dešifrirati sa x_1 ili x_2, tj. $d_K(y)$ ne bi bilo definirano. U skladu s tim imamo da ako je $\boldsymbol{P}$=$\boldsymbol{C}$, onda su funkcije e_K permutacije.

1.2.1 Klasifikacija kriptosustava

Kriptosustavi se obično klasificiraju s obzirom na sljedeća tri kriterija:

1. Tip operacija koje koriste pri šifriranju:

   (a) Supstitucijske šifre - svaki element otvorenog teksta se zamjenjuje s nekim drugim elementom, npr. $TAJNA \to XIWOI$.
   Najčešći primjer ovakve šifre je tzv. *Cezarova šifra* koju je koristio znameniti rimski vojskovođa i državnik Gaj Julije Cezar u komunikaciji sa svojim prijateljima. Šifriranje se radilo tako da su se slova otvorenog teksta zamjenjivala onim slovima koja su se nalazila tri mjesta dalje od njih u alfabetu $(A \to D, B \to E, \dots)$. Pretpostavljamo da se alfabet ciklički nastavlja, tj. da nakon zadnjeg slova Z, ponovno dolaze A, B, C.
   Danas se Cezarovom šifrom nazivaju i šifre istog oblika s pomakom različitim od 3 te možemo uvesti sljedeću definiciju:

   Definicija 1.2. (Cezarova šifra) *Neka je* $\boldsymbol{P} = \boldsymbol{C} = \boldsymbol{K} = \mathbb{Z}_{26}$.
   Za $0 \leqslant K \leqslant 25$ *definiramo:*

   $$e_K(x) = x + K \pmod{26},$$

   $$d_K(y) = y - K \pmod{26}.$$

   Šifra je definirana nad $\mathbb{Z}_{26}$ budući da koristimo 26 slova.

   (b) Transpozicijske šifre - elementi otvorenog teksta permutiraju, npr. $TAJNA \to JANAT$ gdje smo napravili transpoziciju.

   Postoje i sustavi koji kombiniraju obje navedene metode.

2. Način na koji se obrađuje otvoreni tekst:

   (a) Blokovne šifre - obrađuje se blok po blok elemenata otvorenog teksta koristeći jedan te isti ključ K.

   (b) Protočne šifre - obrađuje se element po element otvorenog teksta koristeći pri tome niz ključeva (engl. *keystream*) koji se paralelno generira.

3. Tajnost i javnost ključeva:

   (a) Simetrični (konvencionalni) kriptosustavi - ključ za dešifiranje se može izračunati poznavajući ključ za šifriranje i obratno. Ti ključevi su najčešće identični. Sigurnost leži u tajnosti ključa.

   (b) Kriptosustavi s javnim ključem - ključ za dešifriranje se ne može izračunati iz ključa za šifriranje, barem ne u nekom razumnom vremenu. Kod takvih kriptosustava ključ za šifriranje je *javni*; bilo tko može šifrirati poruku pomoću njega, ali poruku može dešifrirati samo osoba koja ima odgovarajući ključ za dešifriranje - *privatni (tajni) ključ*. 1976. godine su Whitfield Diffie i Martin Hellman prvi javno iznjeli ideju javnoga ključa.

Auguste Kerckhoffs je 1883. godine napisao knjigu *La Cryptographie Millitare* u kojoj je iznio šest osnovnih zahtjeva za kriptografiju:

1. Šifirani tekst treba biti neprobojan u praksi.

2. Sustav šifriranja treba biti prikladan za korisnike.

3. Ključ treba biti lako pamtljiv i promjenjiv.

4. Šifirani tekst treba biti prenosiv telegrafom.

5. Uređaj za šifriranje treba biti lako prenosiv.

6. Uređaj za šifriranje treba biti moguće relativno lako koristiti.

U nastavku ćemo pojasniti Vigenèreovu i Playfairovu šifru koje su primjeri polialfabetskih šifri jer se svako slovo otvorenog teksta može preslikati u jedno od m mogućih slova u ovisnosti o svome položaju unutar otvorenoga teksta, pri čemu je m duljina ključa.

Poglavlje 2

Vigenèreova šifra

U ovom ćemo poglavlju ukratko reći nešto o Giovanu Battisti Bellasu i Blaise de Vigenèreu, osobama koje su doprinijele razvoju Vigenèreove šifre, te pojasniti i ilustrirati na primjerima kako funkcionira ova šifra.

2.1 Razvoj šifre

(a) Giovan Battista Bellaso

(b) Blaise de Vigenère

Slika 2.1: Autori

Giovan Battista Bellaso (1505.-nepoznato) je bio talijanski kriptograf. Prvi puta spominje ovu šifru 1553. godine u svojoj knjizi *La cifra del. Sig. Giovan Battista Bellaso*, ali je u 19. stoljeću otkriće ove šifre pogrešno pripisano Vigenèreu te je po njemu i dobila ime.

Blaise de Vigenère (1523.-1596.) je bio francuski diplomat i kriptograf. Rođen je u mjestu Saint-Pourçain u Francuskoj. Sa 17 godina je stupio u diplomatsku službu, gdje je ukupno proveo 30 godina. Neko vrijeme proveo je u Wormskom saboru te u Rimu gdje je dolazio u kontakt s knjigama o kriptografiji i kriptolozima. U mirovini je napisao preko 20 knjiga, uključujući i *Traktat o kometama (Traicté de Cometes)* 1580. i *Traktat o šiframa (Traicté de Chiffres)* 1585., gdje je opisao šifru koja je kasnije po njemu i dobila ime kao i šifru s "autoključem" koju je sam otkrio. Ta šifra je bila prva šifra dotada koja se nije mogla lako razbiti.

2.2 Vigenèreova šifra

Vigenèreova šifra je jedan od najpopularnijih kriptosustava u povijesti. U širokoj uporabi je bila još tijekom Američke revolucije, krajem 18. stoljeća, a korištena je i u Američkom građanskom ratu. Šifra je dosta dugo smatrana "neprobojnom", čak je 1917. u uglednom časopisu *Scientific American* objavljeno da je ovu šifru nemoguće razbiti te je s vremenom stekla opis "le chiffre indéchiffrable", francuski za neprobojnu šifru. No, to nije bilo točno jer je još 1854. godine grof Charles Babbage razbio ovu šifru, ali nije objavio svoje rezultate, dok je 1863. godine Kasiski detaljno opisao metodu za razbijanje ove šifre.
Sustav koji se danas naziva Vigenèreova šifra definiran je na sljedeći način:

Definicija 2.1. (Vigenèreova šifra) *Neka je m fiksan prirodan broj. Definiramo $\boldsymbol{P} = \boldsymbol{C} = \boldsymbol{K} = (\mathbb{Z}_{26})^m$. Za ključ $K = (k_1, k_2, \ldots, k_m)$ definiramo:*

$$e_K(x_1, x_2, \ldots, x_m) = (x_1 +_{26} k_1, x_2 +_{26} k_2, \ldots, x_m +_{26} k_m),$$

$$d_K(y_1, y_2, \ldots, y_m) = (y_1 -_{26} k_1, y_2 -_{26} k_2, \ldots, y_m -_{26} k_m).$$

Dakle, slova otvorenog teksta pomičemo za $k_1, k_2, ..., k_m$ mjesta, u ovisnosti o tome na kojem se mjestu u otvorenom tekstu nalaze. Preciznije, pomak ovisi o ostatku koji dobijemo kada poziciju slova podijelimo s duljinom ključa m.
Kod ove šifre osnovni su elementi otvorenog teksta i šifrata "blokovi" koji se sastoje od m slova. No, šifriranje se zapravo provodi "slovo po slovo" pa ovdje nije nužno nadopuniti zadnji blok ako broj slova u otvorenom tekstu nije dijeljiv s m.
Ukoliko se alfabet otvorenog teksta i alfabet šifrata sastoje od slova engleske abecede, koristimo sljedeću tablicu numeričkih reprezentanata tih slova:

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25

Tablica 2.1: Numerički reprezentanti slova engleske abecede

Primjer 2.1. (Vigenèreova šifra) *Neka je m=5 i neka je ključa riječ SIFRA. Njezin numerički ekvivalent je ključ* $K = (18, 8, 5, 17, 0)$.
Pretpostavimo da je otvoreni tekst KRIPTOLOGIJA. Numerički ekvivalent otvorenog teksta je $(10, 17, 8, 15, 19, 14, 11, 14, 6, 8, 9, 0)$. *Šifriranje se provodi na sljedeći način:*

	18	*8*	*5*	*17*	*0*	*18*	*8*	*5*	*17*	*0*	*18*	*8*
$+_{26}$	*10*	*17*	*8*	*15*	*19*	*14*	*11*	*14*	*6*	*8*	*9*	*0*
	2	*25*	*13*	*6*	*19*	*6*	*19*	*19*	*23*	*8*	*1*	*8*

Dakle, šifrat je ovdje CZNGTGTTXIBI. Uočimo da se prvo slovo O preslikalo u G, a drugo u T.

Primjer 2.1. može se ilustrirati i na sljedeći način:

ključ	*S*	*I*	*F*	*R*	*A*	*S*	*I*	*F*	*R*	*A*	*S*	*I*
otvoreni tekst	*K*	*R*	*I*	*P*	*T*	*O*	*L*	*O*	*G*	*I*	*J*	*A*
šifrat	*C*	*Z*	*N*	*G*	*T*	*G*	*T*	*T*	*X*	*I*	*B*	*I*

Vidimo da se ovdje ključ ponavlja u nedogled te, s obzirom na način na koji obrađuje otvoreni tekst, ovu šifru možemo shvatiti kao primjer blokovne šifre.
No, postoje i druge varijante Vigenèreove šifre. Jedna takva je ona s *autoključem*, koja je ujedno i sigurnija od originalne varijante. U njoj otvoreni tekst generira ključ. Naime, originalni ključ se koristi samo za šifriranje prvog bloka od m slova, dok se za šifriranje daljnjih blokova koristi prethodni blok otvorenog teksta, zbog čega ona spada u protočne šifre.

Primjer 2.2. (Vigenèreova šifra s autoključem) *U šifriranju se koristi tzv. Vigenèreov kvadrat (Slika 2.2). Ako npr. slovo K treba šifrirati ključem S, onda pogledamo stupac koji počinje s K i redak koji počinje sa S. U presjeku se nalazi šifrat C. Postupak nastavljamo analogno te dobivamo:*

ključ	*S*	*I*	*F*	*R*	*A*	*K*	*R*	*I*	*P*	*T*	*O*	*L*
otvoreni tekst	*K*	*R*	*I*	*P*	*T*	*O*	*L*	*O*	*G*	*I*	*J*	*A*
šifrat	*C*	*Z*	*N*	*G*	*T*	*Y*	*C*	*W*	*V*	*B*	*X*	*L*

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
A	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
B	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A
C	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B
D	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C
E	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D
F	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E
G	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F
H	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G
I	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H
J	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I
K	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J
L	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K
M	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L
N	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M
O	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N
P	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
Q	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
R	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
S	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
T	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S
U	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T
V	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U
W	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V
X	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W
Y	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X
Z	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y

Slika 2.2: Vigenèreov kvadrat

2.3 Kriptoanaliza

Cilj kriptoanalize je određivanje duljine ključne riječi i slova koje ta ključna riječ sadrži.

2.3.1 Određivanje duljine ključne riječi

U nastavku ćemo pojasniti dvije metode za određivanje duljine ključne riječi koje se najčešće koriste:

Kasiskijev test

Kasiskijev test je metoda koju je 1863. godine uveo Friedrich Kasiski, a otprilike u isto vrijeme koristio ju je i Charles Babbage.
Metoda se zasniva na činjenici da će dva identična segmenta otvorenog teksta biti šifrirana na isti način ukoliko se njihove početne pozicije razlikuju za neki višekratnik od m, gdje je m duljina ključa. Obratno, ako uočimo dva identična segmenta u šifratu duljine barem 3, tada je vrlo vjerojatno da oni odgovaraju identičnim segmentima u otvorenom tekstu. Podudarnost odsječaka duljine 2 može lako biti slučajna, ali kod odsječaka veće duljine to je manje vjerojatno.
U Kasiskijevom testu u šifratu tražimo nizove identičnih segmenata duljine barem 3 te, ako takvi postoje, zabilježimo udaljenosti između njihovih početnih položaja. Pomoću tih udaljenosti može se ustanoviti duljina ključa na sljedeći način: ako dobijemo uda-

ljenosti $d_1, d_2, \ldots$, onda je razumno za pretpostaviti da m dijeli većinu d_i-ova. Uzmimo da se npr. prvi niz ponavlja svakih 20 slova. Za taj se niz označe brojevi 2, 4, 5, 10 i 20, tj. djelitelji. Taj se postupak ponovi za svaki niz. Konačno, duljinu ključa određuje onaj broj u tablici koji se najviše puta zabilježio.

Primjer 2.3. *Odrediti moguću duljinu ključa u sljedećem šifratu koristeći Kasiskijev test:*

GSIQITUKQIEAOHRVUGLTAZGHXUHLPJMRTTNQRBZIAVBTG
QTBYMYAIVOMZTAIXJBTEDEWVQWADVWGOOKNQNTCIPEGPY
BOKUSECNWELLCPZUMIVWFUIJMYATUEXISLMZTNPGUJHTM
ERXJSYSIVWABGVWFDTZILNTIEDEFJMFAMPNQZBRSDIZPR
MLGVKFEDZXMVXVQMJXWSLEEQRMAEPRUJXIMFNT

Uočavamo pojavu nekoliko trigrama koji se dva puta pojavljuju u šifratu. To su MYA s početkom na pozicijama 50 i 115 $(115 - 50 = 65 = 5 \cdot 13)$*, MZT s početkom na pozicijama 56 i 125* $(125 - 56 = 69 = 3 \cdot 23)$*, EDE s početkom na pozicijama 65 i 160* $(160-65 = 95 = 5\cdot 19)$*, IVW s početkom na pozicijama 108 i 143* $(143-108 = 35 = 5\cdot 7)$ *i VWF s početkom na pozicijama 109 i 149* $(149-109 = 40 = 5\cdot 8)$*. Uočavamo kako se kao najvjerojatnija duljina riječi nameće broj m=5 koji dijeli sve osim jedne od razlika početnih pozicija ponovljenih trigrama.*

Indeks koincidencije

William Friedman je uveo ovaj pojam prvi puta 1920. godine u svojoj knjizi *Indeks koincidencije i njegove primjene u kriptografiji* koja se smatra jednom od važnijih publikacija u povijesti kriptologije.

Definicija 2.2. (Indeks koincidencije) *Neka je $x = x_1x_2\ldots x_n$ niz od n slova. Indeks koincidencije od x, u oznaci $I_c(x)$, definira se kao vjerojatnost da su dva slučajna elementa iz x jednaka. Neka su $f_0, f_1, \ldots, f_{25}$ redom (apsolutne) frekvencije od A, B, C,..., Z u x. Dva elementa iz x možemo odabrati na $\frac{n(n-1)}{2}$ načina, a za svaki $i = 0, 1, \ldots, 25$ postoji $\frac{f_i(f_i-1)}{2}$ načina odabira dvaput i-tog slova. Stoga vrijedi formula:*

$$I_c(x) = \sum_i \frac{f_i(f_i-1)}{n(n-1)}.$$

Pretpostavimo sada da x predstavlja tekst na hrvatskom jeziku. Očekivane vjerojatnosti pojavljivanja slova A, B, $\ldots$, Z u hrvatskom jeziku označit ćemo redom s $p_0, p_1, \ldots, p_{25}$. Ako je n dovoljno velik, za očekivati je da će vrijediti:

$$I_c(x) \approx \sum_i p_i^2 \approx 0.064.$$

Odnosno, vjerojatnost da su oba slova A je $p_0 \approx 0.115^2$, da su oba B je $p_1 \approx 0.115^2$, itd. Isti zaključak vrijedi i ukoliko je x šifrat dobiven iz otvorenog teksta na hrvatskom jeziku pomoću neke monoalfabetske šifre, kod koje svakom slovu otvorenog teksta odgovara jedinstveno slovo šifrata.
Pretpostavimo sada da imamo šifrat $y = y_1 y_2 \dots y_n$ koji je dobiven Vigenèreovom šifrom. Rastavimo y na m podnizova $z = z_1 z_2 \dots z_m$ tako da y napišemo po stupcima u matricu dimenzije $m \times (n/m)$. Ako m ne dijeli n, možemo nadopuniti y na proizvoljan način ili promatrati "krnju matricu" s nepotpunim zadnjim retkom. Retci ove matrice su upravo traženi podnizovi $z = z_1 z_2 \dots z_m$.
Ako je m jednak duljini ključne riječi, onda su elementi istog retka matrice šifrirani pomoću istog ključa. Na primjer, prvi redak sadrži prvo, $(m+1)$-vo, $(2m+1)$-vo, ... slovo šifrata i sva su ta slova šifrirana pomoću k_1. Zato bi svi indeksi koincidencije $I_c(z_i)$ trebali biti približno jednaki 0.064.
Ako m nije jednak duljini ključne riječi, onda će z_i-ovi izgledati više-manje kao slučajni nizovi slova, budući da su dobiveni pomacima pomoću različitih slova ključa. Primijetimo da za potpuno slučajni niz imamo:

$$I_c \approx 26 \cdot \frac{1}{26^2} = \frac{1}{26} \approx 0.038.$$

Vrijednosti $K_p = 0.064$ i $K_r = 0.038$ (p = plaintext, r = random) su dovoljno daleke jedna od druge tako da ćemo najčešće na ovaj način moći odrediti točnu duljinu ključne riječi - ili potvrditi pretpostavku dobivenu pomoću Kasiskijeve metode.
Spomenimo da se K_p razlikuje ovisno o jeziku te je tako u engleskom jeziku jednak 0.065, u njemačkom 0.076, u francuskom 0.078, u talijanskom 0.074, a u španjolskom 0.078. No, za primjenu same metode nije nužno znati na kojem je jeziku pisan otvoreni tekst, jedino je bitno da se za jezik na kojem je pisan otvoreni tekst veličina K_p značajno razlikuje od 0.038.

Primjer 2.4. *Odrediti moguću duljinu ključa u šifratu navedenom u Primjeru 2.3.: Za m=1 je $I_c = 0.040$, za m=2 su indeksi 0.037 i 0.040, za m=3 su indeksi 0.038, 0.048 i 0.038, za m=4 su 0.036, 0.038, 0.044, 0.036, dok za m=5, dobivamo indekse 0.057, 0.052, 0.066, 0.076, 0.066. Sada smo već prilično sigurni da je duljina ključne riječi jednaka 5.*

2.3.2 Određivanje ključne riječi

Sada kada znamo duljinu ključne riječi, pitanje je kako ju odrediti. Za to nam može pomoći *međusobni indeks koincidencije dvaju nizova*.

Definicija 2.3. (Međusobni indeks koincidencije dvaju nizova) *Neka su* $x = x_1x_2\ldots x_n$ *i* $y = y_1y_2\ldots y_{n'}$ *dva niza od n, odnosno n' slova. Međusobni indeks koincidencije od x i y, u oznaci* $MI_c(x, y)$*, definira se kao vjerojatnost da je slučajni element od x jednak slučajnom elementu od y. Ako frekvencije od A,B,...,Z u x i y označimo s* $f_0, f_1, \ldots, f_{25}$*, odnosno* $f'_0, f'_1, \ldots, f'_{25}$*, onda je*

$$MI_c = \sum_i \frac{f_i \cdot f'_i}{n \cdot n'}.$$

Neka je sada m duljina ključne riječi te neka su $z_1, z_2, \ldots, z_m$ podnizovi dobiveni kao prije. Pretpostavimo da je $K = (k_1, k_2, \ldots, k_m)$ ključna riječ. Pokušat ćemo ocijeniti indeks $MI_c(z_i, z_j)$. Promotrimo proizvoljno slovo u podnizovima z_i i z_j. Procijenimo vjerojatnost da su oba slova jednaka npr. A. Vjerojatnost da pomakom za k_i dobijemo slovo A približno je jednaka vjerojatnosti da se u hrvatskom jeziku pojavljuje slovo čiji je numerički ekvivalent $-k_i \pmod{26}$. Dakle, vjerojatnost da su oba ova slova jednaka A približno je jednaka: $p_{-k_i} \cdot p_{-k_j}$, da su oba slova B približno je jednaka: $p_{1-k_i} \cdot p_{1-k_j}$, itd. Dakle, imamo ocjenu:

$$MI_c(z_i, z_j) \approx \sum_h p_{h-k_i} \cdot p_{h-k_j} = \sum_h p_h \cdot p_{h+k_i-k_j},$$

tj. pomakom indeksa suma se ne mijenja. Uočimo da ova ocjena ovisi samo o razlici $(k_i - k_j) \pmod{26}$ koju ćemo zvati relativni pomak od z_i i z_j. Također vrijedi

$$\sum_h p_h \cdot p_{h+q} = \sum_h p_h \cdot p_{h-q},$$

što znači da za pomak q dobivamo istu ocjenu kao i za pomak 26-q te je stoga dovoljno promatrati pomake između 0 i 13, kao u sljedećoj tablici:

relativni pomak	očekivana vrijednost	relativni pomak	očekivana vrijednost
0	0.064	7	0.033
1	0.039	8	0.040
2	0.031	9	0.042
3	0.031	10	0.036
4	0.044	11	0.036
5	0.040	12	0.036
6	0.039	13	0.039

Tablica 2.2: Relativni pomak i očekivana vrijednost od MI_c

Važno je za uočiti da je ocjena 0.064 ako je pomak jednak 0, a ako je pomak različit od 0, onda su ocjene između 0.031 i 0.044, tj. bitno manje; što nam može pomoći za određivanje vrijednosti $q = k_i - k_j$.
Pretpostavimo da smo fiksirali z_i i promotrim efekt šifriranja z_j sa slovima *A, B, C, ..., Z* (tj. pomakom za 0, 1, 2, ..., 25 mjesta). Tako dobiveni nizove označimo sa $z_j^0, z_j^1, \ldots$. Za g= 0,1,...,25 izračunamo indeks $MI_c(z_i, z_j^g)$ po formuli:

$$MI_c(x, y^g) = \sum_i \frac{f_i \cdot \ f'_{i-g}}{n \cdot n'}.$$

Za $g = q \pmod{26}$, MI_c bi trebao biti blizu 0.064, a za $g \neq q \pmod{26}$ bi trebao uglavnom varirati između 0.031 i 0.044. Na ovaj način se mogu utvrditi relativni pomaci bilo koja dva podniza z_i i z_j. Nakon toga, ostaje nam samo 26 mogućih ključnih riječi koje možemo ispitati jednu po jednu.
No, ukoliko nam je poznato na kojem je jeziku pisan otvoreni tekst, malom modifikacijom ove metode se može doći efikasnije do ključne riječi. Umjesto međusobnog indeksa koincidencije nizova z_i z_j^g, računat ćemo $MI_c(x, z_j^g)$, gdje je x niz koji odgovara tipičnom tekstu na jeziku otvorenog teksta. Pretpostavimo da nam je poznato da je otvoreni tekst pisan na hrvatskom jeziku. To znači da su relativne frekvencije $\frac{f_i}{n}$ približno jednake p_i, pa je:

$$MI_c(x, z_j^g) \approx \sum_i p_i \frac{f'_{i-g}}{n'}.$$

Očekujemo da je $MI_c(x, z_j^g) \approx 0.064$, a u protivnom $MI_c(x, z_j^g) < 0.045$. Prema tome, da bismo odredili j-to slovo k_j ključne riječi, postupamo na sljedeći način: za $0 \leq g \leq 25$ izračunamo $M_g = \sum_i p_i \frac{f'_{i-g}}{n'}$. Odredimo h takav da je $M_h = max M_g : 0 \leq g \leq 25$ te stavimo $k_j = -h \pmod{26}$.

Primjer 2.5. *Odredimo ključnu riječ iz Primjera 2.3. pomoću međusobnog indeksa koincidencije. Već smo zaključili da je m=5. Za j=1,2,3,4,5 izračunajmo vrijednosti* $M_0, M_1, \ldots, M_{25}$.
Npr. za j=0 je

$$M_0 = (0.115 \cdot 3 + 0.015 \cdot 1 + 0.028 \cdot 0 + \cdots + 0.023 \cdot 1)/44 \approx 0.0310.$$

Sve vrijednosti prikazane su u Tablici 2.3 te iščitavamo podatke:
Za j=1, imamo h=14, $M_{14} = 0.0636$*, pa je* $k_1 = -14 \pmod{26}$ *= 26-14 = 12;*
Za j=2, imamo h=0, $M_0 = 0.0616$*, pa je* $k_2 = 0$*;*
Za j=3, imamo h=7, $M_7 = 0.0650$*, pa je* $k_3 = 19$*;*

Za j=4, imamo h=19, $M_{19} = 0.0643$*, pa je* $k_4 = 7$*;*
Za j=5, imamo h=22, $M_{22} = 0.0681$*, pa je* $k_5 = 4$*.*

j	Vrijednosti od M_g za g=0,1, ..., 25						
1	0.0310 0.0342 **0.0636** 0.0364	0.0341 0.0400 0.0440 0.0389	0.0425 0.0471 0.0335 0.0457	0.0427 0.0476 0.0304 0.0395	0.0340 0.0290 0.0400 0.0373	0.0386 0.0266 0.0327	0.0439 0.03550 0.0361
2	**0.0616** 0.0302 0.0421 0.0414	0.0421 0.0423 0.0382 0.0410	0.0291 0.0498 0.031 0.0316	0.0271 0.0411 0.0485 0.0254	0.0449 0.0312 0.0375 0.0383	0.0406 0.0316 0.0373	0.0368 0.0432 0.0345
3	0.0407 **0.0650** 0.0271 0.0408	0.0369 0.0355 0.0393 0.0404	0.0388 0.0332 0.0417 0.0261	0.0439 0.0313 0.0363 0.0415	0.0296 0.0500 0.0385 0.0391	0.0336 0.0407 0.0250	0.0375 0.0392 0.0474
4	0.0347 0.0303 0.0385 0.0357	0.0321 0.0267 0.0496 0.0344	0.0335 0.0429 0.0262 0.0382	0.0289 0.0393 0.0330 0.0450	0.0378 0.0464 0.0462 0.0454	0.0510 0.0250 **0.0643**	0.0339 0.0403 0.0396
5	0.0472 0.0346 0.0380 0.0416	0.0419 0.0398 0.0243 **0.0681**	0.0431 0.0396 0.0362 0.0359	0.0333 0.0338 0.0467 0.0330	0.0420 0.0303 0.0537 0.0274	0.0380 0.0394 0.0212	0.0371 0.0418 0.0312

Tablica 2.3: Međusobni indeks koincidencije

Stoga je ključna riječ MATHE, a otvoreni tekst koji smo tražili je:

USPJE HURJE SAVAN JUNEP OZNAT IHSIF ARAMJ ERISE OVIMC ETIRI
MAPOK AZATE LJIMA REDOM KAKOS UOVDJ ENAVE DENIU PORNO SCUPA
ZLJIV IMPOS TUPCI MAANA LIZEI NTUIC IJOMI SRECO MSPOS OBNOS
TDASE ZNABA REMCI TATIJ EZIKO RIGIN ALNOG TEKST AVEOM AJEPO
ZELJN AALIN IJEBI TNA

Tj., s umetnutim razmacima, interpunkcijom i "kvačicama": ***"Uspjeh u rješavanju nepoznatih šifara mjeri se ovim četirima pokazateljima, redom kako su ovdje navedeni: upornošću, pažljivim postupcima analize, intuicijom i srećom. Sposobnost da se zna barem čitati jezik originalnog teksta veoma je poželjna, ali nije bitna."****, što su prve dvije rečenice Udžbenika za rješavanje vojnih šifara autora Parkera Hitta, jednog od najpoznatijih američkih kriptografa iz Prvog svjetskog rata.*

Poglavlje 3

Playfairova šifra

U ovom ćemo poglavlju ukratko reći nešto o Charlesu Wheatstoneu i Lyonu Playfairu, osobama koje su doprinijele razvoju Playfairove šifre, te pojasniti i ilustrirati na primjerima kako funkcionira ova šifra.

3.1 Razvoj šifre

(a) Charles Wheatstone

(b) Lyon Playfair

Slika 3.1: Autori

Sir Charles Wheatstone (1802.-1875.) je bio engleski znanstvenik zaslužan za neke važne izume iz viktorijanskog doba. Rođen je u Barnwoodu, Gloucester. Bavio se izradom i prodajom glazbenih instrumenata, prevodio francusku poeziju, pisao pjesme. Postaje član švedske znanstvene akademije 1836. godine, a 1873. član francuske znans-

tvene akademije. Bavio se brzinom i strujom. Među brojnim postignućima ističu se i stereoskop i telegraf, kao i njegovi doprinosi u samoj kriptografiji. Šifru o kojoj ćemo u nastavku pričati je 1854. izumio upravo Charles Wheatstone, ali ime je dobila po njegovom prijatelju barunu Playfairu od St. Andrewsa koji ju je popularizirao.
Lyon Playfair (1818.-1898.) je bio švedski znanstvenik i političar. Radio je kao profesor kemije. Bio je pobornik uporabe plinskog otrova protiv Rusa u Krimskom ratu. Postaje predsjednik Društva kemičara 1855. godine.

3.2 Playfairova šifra

Playfairova šifra primjer je polialfabetske šifre gdje su osnovni elementi otvorenog teksta blokovi slova. No, spretno realizirati tu ideju može biti veliki problem jer već za blokove od dva slova treba opisati šifrate od $26^2 = 676$ elemenata otvorenoga teksta.
Ova šifra je bigramska jer se šifriraju parovi slova i to tako da rezultat ovisi o oba slova. Možemo reći da postoji jedna jednostavna podjela ovakvih šifri ovisno o načinu konstrukcije matrice:

1. Playfairova šifra bez ključa: Algoritam za šifriranje se bazira na 5×5 matrici slova koju konstruiramo tako da redom popunjavamo sva slova abecede. Matrica izgleda ovako:

A	B	C	D	E
F	G	H	IJ	K
L	M	N	O	P
Q	R	S	T	U
V	W	X	Y	Z

Tablica 3.1: Playfairova šifra bez ključa

2. Playfairova šifra s ključem: Algoritam za šifriranje se bazira na 5×5 matrici slova koju konstruiramo koristeći ključnu riječ. Unosimo redom i bez ponavljanja prvo sva slova ključne riječi, a zatim preostala slova abecede.
   Na primjer, ako je ključna riječ *PLAYFAIR*, onda matrica izgleda ovako:

P	L	A	Y	F
IJ	R	B	C	D
E	G	H	K	M
N	O	Q	S	T
U	V	W	X	Z

Tablica 3.2: Playfairova šifra s ključem *PLAYFAIR*

U oba slučaja primjećujemo da se, budući da u tablici imamo po 25 slova, dogovorno neka dva slova poistovijete, a najčešće su to I i J. U slučaju da je otvoreni tekst na hrvatskom jeziku, poistovjećivat ćemo V i W da bismo izbjegli moguće nesporazume kod dešifriranja.
Šifriranje se provodi tako da najprije podijelimo otvoreni tekst na blokove od po dva slova. Pritom moramo paziti da se niti jedan blok ne sastoji od dva jednaka slova te da je duljina teksta parna. To postižemo umetanjem npr. slova X gdje god je to potrebno.
Kod šifriranja bloka od dva slova, ovisno o položaju slova u matrici, mogu se dogoditi tri slučaja:

1. Slova se nalaze u istom retku. Tada ih zamijenimo sa slovima koja se nalaze za jedno mjesto udesno ciklički (tj. iza slova koje je skroz desno dolazi slovo koje je skroz lijevo u istom retku). Npr. u Tablici 3.2: $EH \to GK, ST \to TN, FP \to PL$.

2. Slova se nalaze u istom stupcu. Tada ih zamijenimo sa slovima koja se nalaze za jedno mjesto ispod ciklički (tj. iza slova koje je skroz dolje dolazi slovo koje je skroz gore u istom stupcu). Npr. u Tablici 3.2: $OV \to VL, KY \to SC, PI \to IE$.

3. U protivnom, pogledamo pravokutnik koji određuju ta dva slova te ih zamijenimo s preostala dva vrha tog pravokutnika. Redoslijed je određen tako da najprije dođe ono slovo koje se nalazi u istom retku kao prvo slovo u polaznom bloku. Npr. u Tablici 3.2.: $OC \to SR, RK \to CG, PD \to FI$.

Slijedi nekoliko primjera:

Primjer 3.1. (Playfairova šifra bez ključa) *Neka je otvoreni tekst TVOJA JE TAJNA SKRIVENA. Niz pripadajućih bigrama je TV OJ AJ ET AJ NA SK RI VE NA. Šifrirajmo otvoreni tekst koristeći Tablicu 3.1 :*
Rješenje: $TVOJAJETAJNASKRIVENA \to QYTODFDUDFLCUHTGZALC$

Primjer 3.2. (Playfairova šifra s ključem) *Neka je ključna riječ PLAYFAIR, a otvoreni tekst KRIPTOGRAFIJA. Budući da smo slova I i J poistovijetili, a trebali bi se nalaziti u jednom bloku zajedno, dodat ćemo slovo X između. Niz pripadajućih bigrama je KR IP TO GR AF IX JA. Šifrirajmo otvoreni tekst koristeći Tablicu 3.2 : Rješenje:* $KRIPTOGRAFIXJA \rightarrow GCEINQOGYPCUBP$

3.3 Kriptoanaliza

Dešifriranje je samo inverzni postupak od šifriranja s manjim nedoumicama kada su u pitanju npr. slova I i J koje smo poistovijetili. Tada se ovisno o kontekstu odredi koje je slovo u pitanju.

1. Slova se nalaze u istom retku. Tada ih zamijenimo sa slovima koja se nalaze za jedno mjesto ulijevo ciklički (tj. iza slova koje je skroz lijevo dolazi slovo koje je skroz desno u istom retku). Npr. u Tablici 3.2: $GK \rightarrow EH, TN \rightarrow ST, PL \rightarrow FP$.

2. Slova se nalaze u istom stupcu. Tada ih zamijenimo sa slovima koja se nalaze za jedno mjesto iznad ciklički (tj. iza slova koje je skroz gore dolazi slovo koje je skroz dolje u istom stupcu). Npr. u Tablici 3.2: $VL \rightarrow OV, SC \rightarrow KY, IE \rightarrow PI$.

3. U protivnom, pogledamo pravokutnik koji određuju ta dva slova te ih zamijenimo s preostala dva vrha tog pravokutnika. Redoslijed je određen tako da najprije dođe ono slovo koje se nalazi u istom retku kao prvo slovo u polaznom bloku. Npr. u Tablici 3.2.: $SR \rightarrow OC, CG \rightarrow RK, FI \rightarrow PD$.

Za dešifriranje su bitne frekvencije pojedinih bigrama te ćemo stoga spomenuti najfrekventnije bigrame. U hrvatskom jeziku to su: *JE* 2.7 %, *NA* i *RA* 1.5 %, te *ST, AN, NI, KO, OS, TI, IJ, NO, EN, PR* s frekvencijom 1.0 %. U engleskom jeziku to su: *TH* 3.2 %, *HE* 2.5 %, te *AN, IN, ER, RE, ON, ES, TI, AT* 1.2 %. Zanimljivo je da su svi najfrekventniji bigrami oblika suglasnik-samoglasnik ili samoglasnik-suglasnik, osim bigrama *ST* i *PR*.
Također, često nam "posao" olakšava uvođenje dodatne pretpostavke da nam je poznata jedna riječ koja se pojavljuje u otvorenom tekstu, tj. koristimo metodu "vjerojatne riječi". Metoda se sastoji u tome da se napravi lista riječi ili fraza za koje pretpostavljamo da ih otvoreni tekst sadrži te da u šifratu pronađemo segment čija se struktura podudara sa strukturom vjerojatne riječi. Recimo da je vjerojatna riječ *MATEMATIKA I MATEMATIČARI*. Kod rastavljanja na blokove po dva slova imamo dvije mogućnosti, ovisno o tome gdje se ovaj tekst nalazi unutar poruke:

MA	TE	MA	TI	KA	IM	AT	EM	AT	$IČ$	AR	$I*$
$*M$	AT	EM	AT	IK	AI	MA	TE	MA	TI	$ČA$	RI

Sada u šifratu tražimo segment koji ima neku od ove dvije strukture (po dva ponavljanja s istim ovakvim razmacima). Ovakav napad je često prilično efikasan. U praksi se on sprječava tako da se vjerojatne riječi najprije "kodiraju" (npr. $MATEMATIKA = ABC$, $MATEMATIČARI = ADE$), a tek potom šifriraju.

Primjer 3.3. *Dekriptirati sljedeći šifrat dobiven Playfairovom šifrom s ključem (V=W). Otvoreni tekst je na hrvatskom jeziku.*

CK	FL	ET	IJ	KS	VI	XG	IE	QO	SA
GA	PL	TE	AU	KH	CA	ET	AF	CK	TO
IV	OV	OI	AD	BS	HM	HA	OJ	AF	VU
OL	HN	TY	LS	AB	PJ	OL	PJ	AI	LO
AJ	HT	CS	VI	OJ	VC	UI	VI	XG	IE
ZC	AI	BS	HM	IE	AJ	SJ	NT	SJ	AF
OJ	VC	ZA	PJ	JT	HT	ET	OJ	FJ	

Rješenje: Ovaj šifrat je prekratak da bismo ga dekriptirali samo analizom frekvencije bigrama. Za potrebe našeg primjera ćemo pretpostaviti da otvoreni tekst sadrži riječ MINISTARSTVO (u nekom padežu). Struktura ove riječi je: MI NI ST AR ST VO. *Identičnu strukturu uočavamo u 4. retku šifrata i krećemo s pretpostavkom da je* LS AB PJ OL PJ *šifrat od* MI NI ST AR ST. *To je u skladu s analizom frekvencije bigrama:* OJ *(4),* PJ, AF, IE, ET, VI *(3) jer visoko frekventni bigram u hrvatskom jeziku* ST *odgovaa visoko frekventnom bigramu* PJ *u šifratu. Dakle, imamo:* $LS \to MI, AB \to NI, PJ \to ST, OL \to AR$.*
Pokušajmo krenuti u popunjavanje kvadrata za šifriranje. Krenimo prvo s prvom, trećom i četvrtom supstitucijom. Uz pretpostavku da se slova u tim supstitucijama ne nalaze u istom retku ili stupcu, dobivamo do na premutaciju redaka ili stupaca ili njihovo spajanje sljedeću konfiguraciju:*

	A		O
	L	M	R
T		J	
P	I	S	

Za očekivati je da OJ *odgovara nekom visoko frekventnom bigramu (osim* ST *i* RA*), a to su* JE, NA *i* AN. *Ako je* $JE \to OJ$, *onda su* E, J, O *susjedna slova u istom retku ili istom stupcu pa taj slučaj možemo eliminirati. To znači da* O, J, N, A *ili čine pravokutnik ili se nalaze u istom retku ili stupcu. Jedina je mogućnost da retke s* A *i* O *te* T *i* J *spojimo pa tom retku još pridodamo* N. *Taj redak izgleda:* `_AJ_NO_` *ili* `_NJ_AO_`,

s tim da još treba na neko prazno mjesto ubaciti slovo T. No, ovih 5 slova su toliko "razbacana" u abecedi da je jasno daje ovaj redak dio ključa. Zaključujemo da je taj redak TAJNO te da je to prvi redak. Također, vrlo je vjerojatno da je segment PIS dio ključa te ga pomaknimo u drugi redak. Tako dobivamo:

$$\begin{array}{ccccc} T & A & J & N & O \\ P & I & S & & \\ & L & M & & R \end{array}$$

Sada se vrlo vjerojatnom čini pretpostavka da je ključna riječ TAJNOPIS. Popunimo kvadrat i provjerimo pretpostavku:

$$\begin{array}{ccccc} T & A & J & N & O \\ P & I & S & B & C \\ D & E & F & G & H \\ K & L & M & Q & R \\ U & V & X & Y & Z \end{array}$$

Dešifriranjem slijedi:

$$\begin{array}{cccccccccc} PR & EM & DA & SA & MP & LA & YF & AI & RN & IJ \\ EN & IK & AD & TV & RD & IO & DA & JE & PR & ON \\ AL & AZ & AC & TE & SI & FR & EO & NA & JE & UZ \\ AR & GO & NU & MI & NI & ST & AR & ST & VA & RA \\ TA & DO & BI & LA & NA & ZI & VP & LA & YF & AI \\ RO & VA & SI & FR & AI & TA & JX & JO & JX & JE \\ NA & ZI & VO & ST & AO & DO & DA & NA & SX & \end{array}$$

Umetanjem razmaka, interpunkcije i "kvačica" te eliminacijom X -eva, koje smo ranije umetnuli kako u istom bloku ne bi imali dva jednaka slova i kako bi duljina teksta bila parna, dobivamo: ***"Premda sam Playfair nije nikad tvrdio da je pronalazač te šifre, ona je u žargonu Ministarstva rata dobila naziv Playfairova šifra i taj joj je naziv ostao do danas."***

U Drugom svjetskom ratu u redovima njemačke vojske se koristila i dvostruka Playfairova šifra. Ona koristi dva ključa i dvije 5×5 matrice. Šifriranje je jednako kao kod obične Playfairove šifre, s time da se prvo slovo u bigramu smješta u lijevu, a drugo slovo bigrama u desnu matricu. Tako dobijemo pravokutnik za koji vrijedi već spomenuti način šifriranja. Dva slova u paru očito ne mogu biti u istom stupcu, ali i dalje mogu biti u istom retku, što također znamo šifrirati.
Playfairova šifra ima nekoliko prednosti u usporedbi sa supstitucijskom šifrom. Budući je šifra bigramska, u šifratu se gube jednoslovne riječi što utječe na frekvencije. Nadalje,

bigramsko šifriranje smanjuje na polovicu broj elemenata dostupnih analizi frekvencije. S druge strane, broj bigrama je puno veći od broja individualnih slova (26 slova - 676 bigrama), dok su frekvencije najčešćih bigrama puno ujednačenije od frekvencija najčešćih slova. Zbog toga je ova šifra dugo smatrana sigurnom te je bila standardna šifra u britanskoj vojsci u Prvom svjetskom ratu, a korištena je i za šifriranje poruka manje važnosti u američkoj vojsci u Drugom svjetskom ratu.

Nedostatak je što kod duljih tekstova postaje nesigurna jer se može iskoristiti analiza frekvencija bigrama. Poznato je da i kod ove šifre dio strukture ostaje sačuvan. Naime, slova u šifratu nisu jednoliko raspoređena. U otvorenom tekstu na engleskom jeziku npr. najfrekventnije slovo ima relativnu frekvenciju $\approx 13\%$, a u šifratu dobivenom ovom šifrom to iznosi $\approx 7\%$, dok kod Vigenèreove šifre imamo $\approx 6\%$. Ti podatci nam mogu pomoći kod određivanja vrste šifre.

No, Playfairova šifra nije preporučljiva za veće blokove. Na primjer, za blokove od 3 slova treba opisati šifrate od $26^3 = 17576$ elemenata otvorenog teksta, a za blokove od 4 slova čak $26^4 = 456976$ i sl. Ipak, postoje ne tako jednostavne i očite generalizacije Playfairova kvadrata u kocku za šifriranje blokova od 3 slova ili u neki četverodimenzionalni objekt za šifriranje blokova od 4 slova (vidi npr. [2] i [6]).

Literatura

[1] C. Brew, *The Playfair Cipher*,
URL: http://www.ling.ohio-state.edu/~cbrew/2008/spring/playfair.pdf

[2] C. Christensen, *Polygraphic Ciphers*,
URL: http://www.nku.edu/~christensen/section%2019%20playfair%20cipher.pdf

[3] A. Dujella, M. Maretić, *Kriptografija*, Element, Zagreb, 2007.

[4] A. Dujella, *Vigenèreova šifra*, math.e, **1**(2004),
URL: http://e.math.hr/vigenere/index.html

[5] A. Galinović, *Povijest kriptografije*,
URL: http://web.zpr.fer.hr/ergonomija/2005/galinovic/index.html

[6] L. Lu, Y. Shang, *An Extended Algorithm based on Playfair Cipher*, National Conference on Information Technology and Computer Science, 2012.
URL: http://www.atlantis-press.com/php/download_paper.php?id=2979.

[7] *Wikipedija*,
URL: https://hr.wikipedia.org/wiki/Blaise_de_Vigen%C3%A8re
URL: https://en.wikipedia.org/wiki/Charles_Wheatstone
URL: https://en.wikipedia.org/wiki/Giovan_Battista_Bellaso
URL: https://hr.wikipedia.org/wiki/Kriptografija
URL: https://hr.wikipedia.org/wiki/Kriptologija
URL: https://en.wikipedia.org/wiki/Lyon_Playfair,_1st_Baron_Playfair
URL: https://en.wikipedia.org/wiki/Vigen%C3%A8re_cipher